

# Subjectiviteit door de jaren heen: conversationalisatie in journalistieke teksten

Kirsten Vis, José Sanders & Wilbert Spooren

*Binnen de traditie van critical discourse analysis heeft Fairclough het concept 'conversationalisatie' geïntroduceerd: in case studies van publiek taalgebruik observeert hij een verschuiving in de tijd, waarbij steeds meer informele kenmerken uit conversatie gebruikt worden. In navolging van Steen (2003) en Pearce (2005) willen we een bijdrage leveren aan de empirische toetsing van deze theorie. We doen dit door het begrip conversationalisatie te operationaliseren in termen van talige uitdrukkingen van subjectiviteit. Subjectiviteit zien we dan (in navolging van Scheibman, 2002 en Bekker, 2006) vooral als een lijst van kenmerken die uitdrukking geven aan de aanwezigheid van de spreker/schrijver. Te denken valt aan persoonlijke voornaamwoorden (eerste en tweede persoon enkel- en meervoud), modale werkwoorden en bijwoorden van graad. Voor het onderzoek hebben we drie subcorpora ingericht: krantenteksten uit 1950, krantenteksten uit 2002 en conversaties uit 2002. De analyse laat zien dat conversaties, zoals verwacht, veel subjectiever zijn dan krantenteksten. De vraag of krantenteksten subjectiever zijn geworden heeft geen eenduidig antwoord. Overall is dat niet het geval maar de hoeveelheid directe rede is wel enorm toegenomen tussen 1950 en 2002. Binnen de fragmenten die niet in de directe rede staan, is de subjectiviteit eerder afgenomen. We interpreteren dit resultaat als volgt: de subjectiviteit die de conversationalisatietheorie veronderstelt, schuilt vooral in de hoeveelheid geciteerde spraak en in de subjectieve weergave van die spraak. In een vervolgonderzoek, met deelcorpora van elk ca. 1 miljoen woorden, zal deze interpretatie nader onderzocht worden.*

## 1 Inleiding<sup>1</sup>

In verschillende wetenschappelijke disciplines, zoals genreanalyse en *critical discourse analysis*, zijn veranderingen in genres opgemerkt (Bhatia, 2005; Askehave & Swales, 2001). Zo stelt Bhatia dat in genres als academische teksten kenmerken doordringen die typerend voor advertentieteksten genoemd kunnen worden. Binnen het kader van de *critical discourse analysis* heeft Fairclough (1992, 1994) het begrip 'conversationalisatie' geïntroduceerd. Fairclough signaleert in publieke communicatie –zoals journalistiek, politiek, gezondheidszorg en onderwijs– een tendens naar een meer informele en conversationele stijl. Als voorbeeld noemt Fairclough een krantenkop uit de Sunday Mirror: “*Di’s butler bows out ... in sneakers*” [*Di’s butler trekt zich terug ... op gympen*] (Fairclough, 1992, p. 111), waarin conversationeel vocabulaire wordt gebruikt. Ook beschouwt hij het gebruik van grafische middelen (de drie puntjes) als een poging om spontane conversatie na te bootsen. Naast voorbeelden uit dagbladen geeft Fairclough ook voorbeelden uit andere genres, zoals universiteitsbrochures en dokter-patiënt-interacties waarin hij een informele en conversationele stijl aanwijst. Faircloughs conversationalisatietheorie is intuïtief plausibel, maar zijn empirisch bewijs beperkt zich tot dergelijke anekdotische voorbeelden uit hedendaags taalgebruik. Ook in een diachrone behandeling van de conversationalisatietheorie als in Fairclough en Mauranen (1997) blijft de bespreking beperkt tot een analyse van een aantal individuele politieke sprekers (uit Groot-Brittannië en Finland).

Andere onderzoekers hebben wel meer systematische aanwijzingen gevonden voor een verschuiving in publieke communicatie in de richting van conversaties. Zo

vond Steen (2003) in een corpusonderzoek van 120 redactionele commentaren uit *The Times* (85.000 woorden) een toename in geïnvolveerdheid en een daling in narrativiteit. Dit kan geïnterpreteerd worden als een stilistische verandering in de richting van conversaties, die immers typisch geïnvolveerd en persuasief van aard zijn, en weinig markering vertonen voor narrativiteit (althans de narrativiteitskenmerken zoals Biber (1988) die onderscheidt: verleden tijd, voltooid aspect en derde persoon).

Ook Cotter (2003) concludeerde op grond van een diachrone corpusanalyse (100.000 woorden) dat kranten eigenschappen van conversatie overnemen: in regionale kranten uit Californië nam in de periode van 1900 tot 1995 de frequentie van *and* en *but* in zinsinitiële positie – een verschijnsel dat typisch geassocieerd wordt met gesproken taal – toe.

Pearce (2005) vond in een diachrone corpusanalyse (37.000 woorden) van 32 Britse politieke toespraken tussen 1966 en 1997 een toename in conversationalisatie, gereflecteerd in linguïstische kenmerken als de daling van het aantal nominalisaties, een toename van het gebruik van voornaamwoorden in 1e en 2e persoon en het gebruik van *that-complement clauses*.

Tot slot concludeerde Biber (2004) uit de resultaten van een diachroon onderzoek (560 teksten) naar verschillende genres (nieuws, drama, brieven en medische teksten) over de periode van 1650 tot 1990 dat er grote verschuivingen zijn in sociale normen die door de uiteenlopende teksten worden weerspiegeld. Het uitdrukken van meningen, houdingen en emoties gebeurt het meest in conversaties (Biber et al., 1999), maar sprekers zijn –vooral in de laatste 50 jaar– ook in de andere genres steeds meer geneigd hun houding of mening uit te drukken.

De resultaten van deze studies bieden aanknopingspunten om de conversationalisatietheorie te toetsen voor Nederlandstalige publieke communicatie, in het bijzonder kranten. Daarvoor kijken we, net als Biber (2004), Cotter (2003), Pearce (2005) en Steen (2003), naar kenmerken van spontane conversatie. Teneinde het begrip conversationalisatie in te bedden in de taalgebruikstheorie passen we de typering toe zoals geformuleerd door Clark (1996): aanwezigheid van de spreker, zelfexpressie van de spreker en expressie van aandacht van de spreker voor de aanwezigheid van de geadresseerde.

Deze kenmerken worden in de literatuur geschaard onder verschillende maar vergelijkbare labels, zoals ‘subjectiviteit’ (in de zin van Langacker, 1990) en ‘*stance*’. Subjectiviteit definiëren we als de expressie door de spreker van zichzelf en zijn attitudes en opvattingen (Lyons, 1994), terwijl *stance* duidt op het uitdrukken van emoties, houdingen en meningen (Biber et al., 1999). Deze definities liggen dicht bij elkaar en de twee begrippen zijn op basis van de literatuur niet altijd goed te onderscheiden. We bespreken eerst een aantal verschillende invalshoeken van waaruit onderzoek is gedaan naar linguïstische kenmerken van subjectiviteit/*stance*.

Biber (2004) onderzocht de diachrone ontwikkeling van *stance* in verschillende genres door drie fenomenen waarmee *stance* uitgedrukt kan worden, onder de loep te nemen: *stance*-uitdrukkende bijwoorden (*hopefully, certainly*), (semi-)modale werkwoorden (*can, must, got to, be going to*) en *stance complement clause* constructies (*suggest that, possible that, hope to, appear to, likely to*).

Scheibman (2002) analyseerde subjectiviteit in Amerikaans-Engelse conversaties. Haar corpusanalyse resulteerde in een lijst van algemene kenmerken van subjectiviteit, zoals persoonlijke voornaamwoorden (eerste en tweede persoon enkel- en meervoud), modale werkwoorden en bijwoorden van graad. Wiebe (1994) ontwikkelde een model voor automatische identificatie en extractie van opinies en emoties en on-

derscheidde op grond van een grootschalige corpusanalyse diverse linguïstische elementen die subjectiviteit uitdrukken, zoals uitroepen, directe vragen en elementen die een evaluatie of oordeel uitdrukken (*awful, incredibly*) en *evidentials* (*surely, might*).

Bekker (2006) onderzocht in een corpus van Nederlandse krantenteksten het verband tussen subjectiviteit en zinsvolgorde. De potentieel subjectieve elementen die ze daartoe heeft onderscheiden, zijn specifiek toegespitst op sprekerssubjectiviteit: uitdrukkingen van modaliteit, van mentale activiteit, communicatieve handelingen en evaluaties.

Zoals uit dit overzicht duidelijk werd, worden onder de begrippen subjectiviteit en *stance* deels dezelfde linguïstische kenmerken behandeld. Wij beschouwen subjectiviteit en *stance* dan ook als sterk gerelateerde begrippen, waarbij *stance* gezien kan worden als dimensie van subjectiviteit, en wij zien de frequentie van linguïstische kenmerken van subjectiviteit (waaronder die voor *stance*) dan als maat voor conversationalisatie.

Bijzondere aandacht verdient het begrip sprekerssubjectiviteit, geïntroduceerd door Bekker (2006). Hiermee wordt bedoeld op de subjectiviteit van de spreker (journalist), die staat tegenover de subjectiviteit van andere personages (journalistieke bronnen) waarover in de tekst gesproken wordt.

Dit onderscheid is in het onderzoek naar subjectiviteit<sup>2</sup> in kranten van groot belang. Subjectiviteit is zoals gezegd de expressie van de schrijver van zichzelf, met andere woorden, het op de voorgrond plaatsen door de spreker van zichzelf en zijn attitudes en opvattingen. In kranten treden echter naast de schrijver van het artikel, de journalist, ook andere sprekers naar voren; vele verschillende bronnen komen immers aan het woord. De spraak van andere sprekers, of personages, kan op verschillende manieren worden geciteerd. In kranten gebeurt dit over het algemeen op één van de volgende manieren:

- Directe rede

Door middel van aanhalingstekens wordt expliciet aangegeven dat een of meer zinnen uitgesproken zijn door een personage. Het citaat wordt ingebed door een communicatief werkwoord als ‘zeggen’ of een woord als ‘aldus’:

1. ‘Dit is voor ons absoluut niet te controleren’, zegt Sevenstar-directeur Koolhof. (Algemeen Dagblad, 3 mei 2002, katern ‘binnenland’)

- Indirecte rede

In de indirecte rede worden de woorden van het personage weergegeven in een afhankelijk bijzin die wordt ingeleid met een communicatief werkwoord zoals ‘zeggen’ of ‘antwoorden’:

2. Naar aanleiding van enkele vragen van de president van de raad prof. mr. Verzijl, antwoordde Fischer, dat hij het gebeurde thans beurt. (NRC, 6 juli 1950, katern ‘binnenland’)

- Semi-directe rede

Naast de directe en indirecte rede bestaat er nog een tussenvorm, die in de Algemene Nederlandse Spraakkunst (Haeseryn et al. 1997) ‘semi-directe rede’ wordt genoemd. Bij de semi-directe rede wordt pas aan het eind van de zin duidelijk dat de zin geïnterpreteerd moet worden als spraak van een ander personage:

3. Operaties kunnen bij milde gevallen echter erg ingrijpend zijn, vindt Berghmans. (Algemeen Dagblad, 10 mei 2002, katern ‘wetenschap’)

- Deelcitaten

Door middel van aanhalingstekens wordt aangegeven dat een of meer woorden (maar niet een hele zin) is uitgesproken door een personage.

4. Volgens bronnen binnen de bank vindt Smits dat het werken hem ‘onmogelijk is gemaakt’. (Trouw, 3 september 2002, katern ‘economie’)

In de directe rede en de deelcitaten wordt door de journalist expliciet –door middel van aanhalingstekens– aangegeven dat een gedeelte van de tekst niet voor zijn verantwoording komt, maar voor de verantwoording van een bepaalde bron (Sanders & Spooren, 1997). In voorbeeld (1) en (4) markeert de journalist dat de gedeeltes tussen aanhalingstekens geheel voor rekening komen van resp. Sevenstar-directeur Koolhof en Smits, voormalig Rabobank-topman. In de indirecte en semi-directe rede (voorbeeld (2) en (3)) is dit onderscheid echter onduidelijker; in die gevallen komt de verwoording van de spraak van de andere personages minstens voor een deel voor de verwoording van de journalist.

Om de twee stemmen –journalist aan de ene kant en andere personages aan de andere kant– uit elkaar te houden, maken we een onderscheid tussen de indirecte en semi-directe rede aan de ene kant en de directe rede en deelcitaten aan de andere kant. In de eerste groep (indirecte rede, semi-directe rede) treedt de stem van de journalist nadrukkelijk op de voorgrond, net als in tekstdelen waarin geen spraak van een ander weergegeven wordt. In de tweede groep (directe rede, deelcitaten) treedt de stem van andere personages nadrukkelijk op de voorgrond. Dit betekent dat de subjectieve elementen die in de directe rede en deelcitaten worden aangetroffen worden gerekend tot de personagesubjectiviteit van de geciteerde bron, terwijl de subjectieve elementen in de indirecte rede en semi-directe rede gerekend worden tot de sprekersubjectiviteit van de vertellende journalist. In de voorbeelden (1) t/m (4) (hieronder herhaald) betekent dit, dat *thans* in voorbeeld (2) en *echter* en *erg* in voorbeeld (3) gerekend worden tot sprekersubjectiviteit, terwijl *absoluut* in voorbeeld (1) en *onmogelijk* in voorbeeld (4) gerekend worden tot personagesubjectiviteit.

1. ‘Dit is voor ons **absoluut** niet te controleren’, zegt Sevenstar-directeur Koolhof. (Algemeen Dagblad, 3 mei 2002, katern ‘binnenland’)
2. Naar aanleiding van enkele vragen van de president van de raad prof. mr. Verzijl, antwoordde Fischer, dat hij het gebeurde **thans** betreurt. (NRC, 6 juli 1950, katern ‘binnenland’)
3. Operaties kunnen bij milde gevallen **echter erg** ingrijpend zijn, vindt Berghmans. (Algemeen Dagblad, 10 mei 2002, katern ‘wetenschap’)
4. Volgens bronnen binnen de bank vindt Smits dat het werken hem ‘**onmogelijk** is gemaakt’. (Trouw, 3 september 2002, katern ‘economie’)

Nu we deze beide aspecten van subjectiviteit onderscheiden hebben, treedt de vraag naar voren of de conversationalisatietheorie zich tot beide aspecten uitstrekt. Onze

veronderstelling is dat indien de conversationalisatietheorie juist is, niet alleen de spreker van de publieke tekst sterker naar de voorgrond zal zijn gaan treden, maar ook de mate waarin andere sprekers in publieke teksten aan het woord komen. Met andere woorden, we verwachten in de loop van de tijd een toename van het aantal subjectiviteitskenmerken in de vertellerstekst (hypothese 1) en een toename van het aandeel citaten van andere sprekers (hypothese 2).

## 2 Methode

In dit hoofdstuk zullen de werkwijze, het analysemodel, het materiaal en de bewerking van gegevens van dit onderzoek besproken worden. Het doel van dit onderzoek is het toetsen van de conversationalisatietheorie voor Nederlandse kranten. Daartoe worden in een corpusanalyse twee vergelijkingen uitgevoerd: in de eerste plaats worden de frequenties van voorkomen van subjectieve elementen in kranten uit 2002 vergeleken met de frequenties van voorkomen in spontane conversaties uit 2002. Deze vergelijking heeft tot doel vast te stellen hoe subjectief hedendaagse krantenteksten zijn, vergeleken met hedendaagse conversaties. Daarna volgt de diachrone vergelijking. Daarin worden de frequenties van voorkomen van de subjectieve elementen in kranten in 1950 vergeleken met de frequenties van voorkomen in kranten in 2002; eerst voor de hele tekst met de twee stemmen (vertellerstekst van de journalist en de direct geciteerde brontekst van de personages) door elkaar en daarna nogmaals voor alleen de sprekerssubjectiviteit (de vertellerstekst van de journalist zonder direct geciteerde brontekst).

Het materiaal bestaat uit drie deelcorpora: twee deelcorpora bestaande uit oude respectievelijk nieuwe kranten, en één deelcorpus van spontane conversaties. In Tabel 1 worden de deelcorpora toegelicht. De teksten zijn automatisch geannoteerd voor woordsoort- en lemma-informatie door het toevoegen van xml-codes.<sup>3</sup>

*Tabel 1 Het materiaal*

Register	Aantal woorden
Kranten uit 1950 (afkomstig uit NRC, de Telegraaf, Trouw en de Volkskrant; gespreid over katernen: voorpagina, binnenland, buitenland, economie, kunst, opinie, sport, wetenschap)	35.000 woorden (50 teksten)
Kranten uit 2002 (afkomstig uit Algemeen Dagblad, NRC Handelsblad, de Telegraaf, Trouw en de Volkskrant; gespreid over katernen: voorpagina, binnenland, buitenland, economie, kunst, opinie, sport, wetenschap)	55.000 woorden (100 teksten)
Conversaties (afkomstig uit het Corpus Gesproken Nederlands (Oostdijk 2002); complete gesprekken)	50.000 woorden (30 gesprekken)

Voor de analyse van subjectieve elementen is een analysemodel opgesteld.<sup>4</sup> In dit model zijn elementen uit de operationalisering van subjectiviteit van Bekker (2006), Biber (2004), Scheibman (2002) en Wiebe (1994) bijeengebracht. Hierbij is in de eerste plaats gekozen voor de elementen die door verschillende auteurs aangehaald wor-

den en in de tweede plaats voor elementen die zich lenen voor automatische analyse. De lijst is gecompleteerd met gebruikmaking van de beschrijvingen van de grammaticale categorieën van *stance* en modaliteit in de grammatica's van Biber et al. (1999) en Haeseryn et al. (1997). In Tabel 2 wordt een overzicht gegeven van dit analysemodel met bij elk element een vermelding van de bron van dit element en enkele voorbeelden. De bijlage toont een deel van een geannoteerde tekst als voorbeeld.

Tabel 2 *Subjectieve elementen in het analysemodel en voorbeelden*

Subjectieve elementen		Bron	Voorbeelden
Voornaamwoorden (1e en 2e persoon enkel- en meervoud)		S	<i>ik, jij, mijn, onze</i>
Modale elementen	Werkwoorden	Be, Bi1999, Bi2004, H, S, W	<i>kunnen, moe- ten, blijken, schijnen</i>
	Bepalingen van modaliteit	Bijwoorden Be, Bi1999, Bi2004, H, S, W	<i>mogelijk, zeker, eigenlijk, hopelijk</i>
	Modale functies van conjunctief en imperatief	Oordeel- partikels H, W H	<i>dan, nog, nou, al, pas Het ga je goed. Kom hier!</i>
Cognitieve werkwoorden		Be, Bi1999, Bi2004, S	<i>zeggen, denken, hopen, verwachten</i>
Intensifiers	Bijwoorden van graad (met versterkende betekenis)	S, W	<i>nogal, hartstikke, zeer</i>
	Kwantificerende bijwoorden	S, W	<i>bijna, nauwelijks, helemaal</i>
Vragen		W	<i>Hoe nu deze crisis te verklaren?</i>
Uitroepen Deiktische elementen <sup>5</sup>		S, W	<i>Wat mooi! nu, hier, gisteren</i>

Be = Bekker (2006), Bi1999 = Biber et al. (1999), Bi2004 = Biber (2004), H = Haeseryn et al. (1997), S = Scheibman (2002), W = Wiebe (1994)

De drie deelcorpora zijn automatisch geannoteerd voor subjectiviteit. Daarbij zijn alle plaatsen in het corpus waar een subjectief element uit het analysemodel gesignaleerd werd, voorzien van subjectiviteitscodes. De gedeelten in directe rede en de deelcitaten zijn handmatig voorzien van codes, zodat ze uit het corpus geïsoleerd konden worden.

Na de codering van het gehele corpus voor subjectiviteit en directe citaten zijn de frequenties van de verschillende subjectieve elementen in de drie deelcorpora berekend. Voor de oude en nieuwe kranten zijn ook de frequenties berekend van alleen

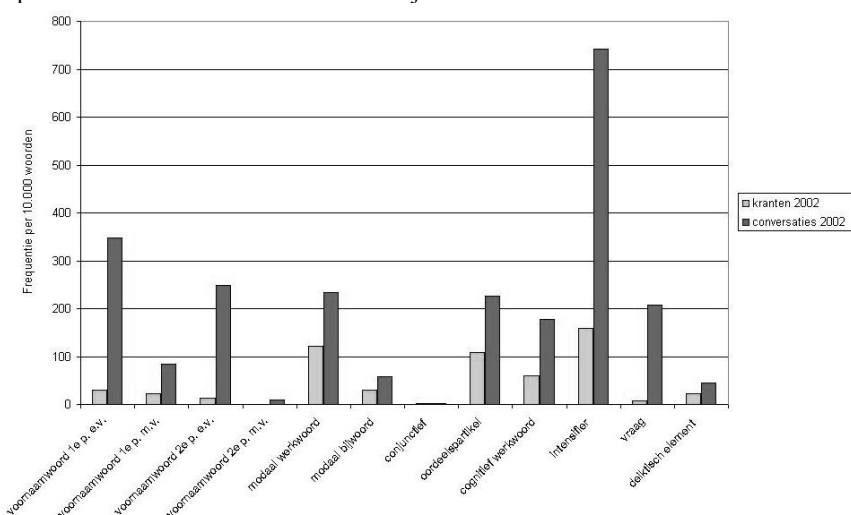
de sprekerssubjectieve elementen, met andere woorden, zonder de personagesubjectieve elementen in de directe citaten mee te rekenen.

De verschillen tussen de verschillende frequenties in de deelcorpora zijn statistisch getoetst met behulp van een loglikelihood-toets<sup>6</sup>. In de volgende paragraaf worden de resultaten van deze berekeningen gepresenteerd. In de presentatie zijn de frequenties omwille van de vergelijkbaarheid genormaliseerd tot 10.000 woorden.

### 3 Resultaten

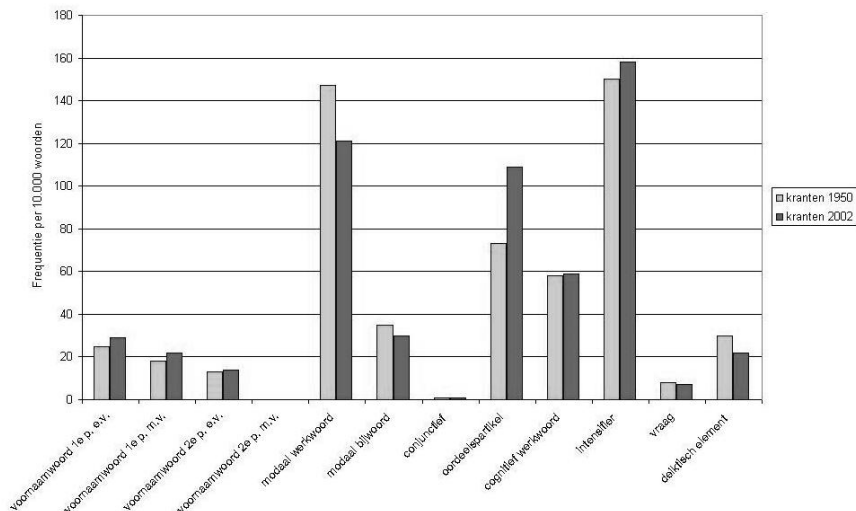
De eerste vergelijking behelst die tussen de frequentie van subjectieve elementen in de kranten uit 2002 en de frequentie van subjectieve elementen in de conversaties uit 2002. In Figuur 1 worden de frequenties, genormaliseerd voor 10.000 woorden, naast elkaar getoond. De resultaten voor beide deelcorpora staan steeds per element bij elkaar; de linkerbalk geeft de frequentie in de kranten weer en de rechterbalk de frequentie in de conversaties.

Figuur 1 geeft een duidelijk beeld: alle elementen komen méér voor in conversaties dan in de kranten. Deze verschillen zijn alle significant (minimale  $G^2$ -waarde 43,56,  $p < .001$ ), met uitzondering van de modale conjunctieven, waarvan de frequentie niet verschilt ( $G^2(1) = 0,07$ ,  $p = 0,79$ ). Dit laat zien dat de onderzochte elementen de potentie hebben om een verschil in subjectiviteit te reflecteren.



Figuur 1 Vergelijking van frequentie van subjectieve elementen in kranten 2002 en conversaties 2002, genormaliseerd voor 10.000 woorden

In hypothese 1 veronderstelden we dat hedendaagse kranten meer subjectieve elementen bevatten dan oude kranten. Dit kan getoetst worden met een diachrone vergelijking van de kranten. Eerst wordt de vergelijking tussen de twee deelcorpora besproken waarbij geen onderscheid gemaakt wordt tussen spreker- en personagesubjectiviteit. In deze vergelijking tellen alle subjectieve elementen in de deelcorpora mee. Deze diachrone vergelijking is weergegeven in Figuur 2.



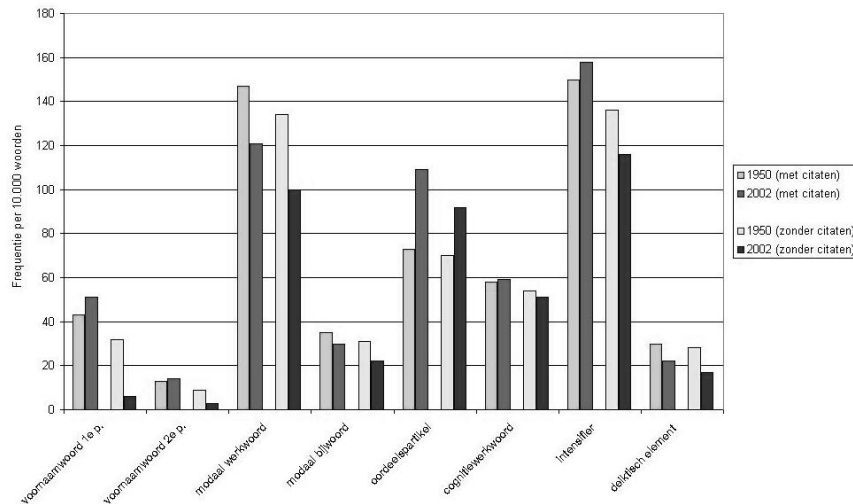
Figuur 2 *Vergelijking van frequentie van subjectieve elementen in kranten 1950 en kranten 2002, genormaliseerd voor 10.000 woorden*

Deze figuur geeft een veel minder duidelijk beeld dan Figuur 1. Er zijn inderdaad elementen waarvan de frequentie toeneemt, maar voor de voornaamwoorden, cognitiewerkwoorden en intensifiers zijn de verschillen niet significant.<sup>7</sup> Slechts drie elementen vertonen een significante verandering: zo nemen de oordeelpartikels in aantal toe (1950: 73,3/10000 woorden, 2002: 108,8/10000 woorden;  $G^2(1)=29,25$ ,  $p < .01$ ). De modale werkwoorden en de deiktische elementen nemen daarentegen significant af (mod. werkwoorden: 1950: 146,9/10000 woorden, 2002: 121,4/10000 woorden,  $G^2(1)=10,4$ ,  $p < .01$ ; deiktische elementen: 1950: 30,2/10000 woorden, 2002: 22,1/10000 woorden,  $G^2(1)=5,45$ ,  $p < .05$ ).

Hypothese 2 gaat over het aandeel van de geciteerde tekst in verhouding tot de vertellertekst. Om deze hypothese te toetsen zijn de aandelen directe rede en deelcitaten in de krant tussen de oude en nieuwe kranten vergeleken. Het blijkt dat dit aandeel tussen 1950 en 2002 is toegenomen: het relatieve aantal woorden in directe rede is significant gestegen (1950: 455 per 10000 woorden, 2002: 1291 per 10000 woorden,  $G^2(1) = 1719,02$ ,  $p < .01$ ). De deelcitaten, die overigens een veel minder frequent verschijnsel vormen, laten het tegenovergestelde effect zien: het relatieve aantal woorden dat als deelcitaat gepresenteerd wordt, is juist significant gedaald (1950: 16,9/10000 woorden, 2002: 11,9/10000 woorden,  $G^2(1)= 37,05$ ,  $p < .01$ ).

Tot slot zijn kranten uit 1950 en 2002 vergeleken zonder personagesubjectieve elementen; met andere woorden, alleen de subjectieve elementen die voor rekening van de journalist komen (sprekerssubjectiviteit) zijn vergeleken. In Figuur 3 zijn de frequenties afgebeeld. Per subjectief element staan vier balken naast elkaar. De eerste twee zijn een herhaling van de frequenties uit Figuur 2; de derde en vierde balk geven de diachrone vergelijking weer van uitsluitend de sprekerssubjectieve elementen.

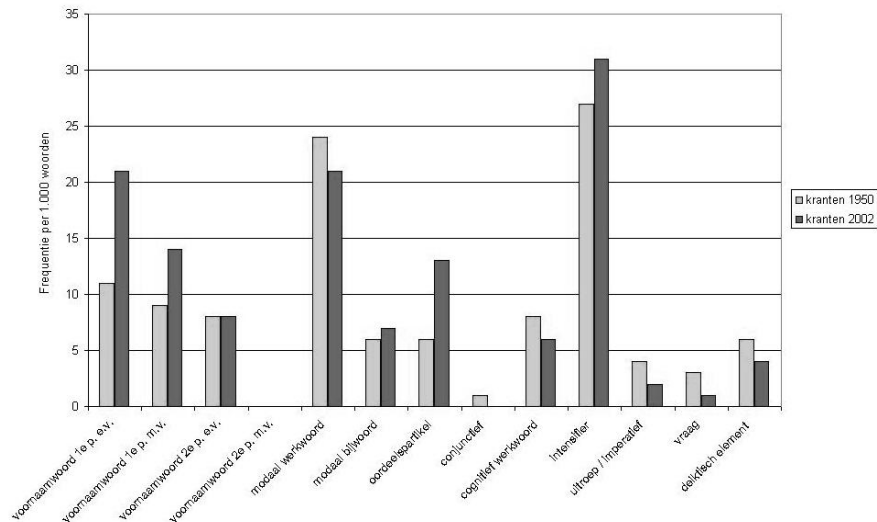




Figuur 3 *Vergelijking van frequentie van subjectieve elementen in kranten 1950 en kranten 2002, met citaten en zonder citaten, genormaliseerd voor 10.000 woorden*

Er is een aantal veranderingen ten opzichte van de eerste diachrone vergelijking (Figuur 2). Een aantal elementen was bij de eerste vergelijking niet-significant gestegen, maar blijkt, nu de broncitaten buiten beschouwing worden gelaten, significant gedaald: voornaamwoorden, modale bijwoorden en intensifiers.<sup>8</sup> Bij twee elementen was de daling al significant, maar dit effect is bij deze vergelijking sterker geworden (modale werkwoorden, deiktische elementen<sup>9</sup>). Het enige element dat in deze vergelijking significant is gestegen, is het oordeelpartikel<sup>10</sup>.

Merk op dat uit de vergelijking in Figuur 3 valt af te leiden dat de frequenties van subjectieve elementen in de directe rede en deelcitaten toenemen met de tijd: immers, als directe rede mede onder ogenschouw wordt genomen (Figuur 2), dan stijgt de frequentie van een aantal elementen of blijft gelijk, terwijl de frequentie van die elementen in de tekst zonder directe rede daalt (Figuur 3). Dit beeld wordt bevestigd in Figuur 4. Met name voor persoonlijke voornaamwoorden 1<sup>e</sup> persoon en oordeelpartikels is de toename significant (1<sup>e</sup> pers. pers.vnw.: 1950: 16,52/1000 w; 2002: 31,78/1000 w;  $G^2(1) = 16,52$ ,  $p < .001$ ; oordeelpartikels: 1950: 5,53/1000 w; 2002: 11,54/1000 w;  $G^2(1) = 6,91$ ,  $p < .01$ ). Voor imperatieven en uitroepen is sprake van een significante afname (1950: 3,22/1000 w; 2002: 1,54/1000 w;  $G^2(1) = 5,73$ ,  $p < .05$ ).



Figuur 4 *Vergelijking subjectieve elementen in de directe rede, genormaliseerd voor 1.000 woorden.*

#### 4 Conclusie en discussie

We hebben aangenomen dat subjectiviteit een aspect van conversationalisatie is. We hebben op grond van de literatuur tekstuele elementen geïdentificeerd die subjectiviteit uitdrukken. Uit de uitgevoerde analyses blijkt dat, zoals verwacht, conversaties meer van dergelijke subjectieve elementen vertonen dan krantenteksten. Dat ondersteunt onze aanname dat het analysemodel geschikt is om de sprekerssubjectiviteit te onderzoeken. De resultaten van de diachrone vergelijkingen wijzen niet op een duidelijke toename van de subjectieve elementen, integendeel zelfs: frequenties van subjectieve elementen nemen, wanneer directe rede en deelcitaten uitgesloten worden, allemaal af. De enige uitzondering hierop zijn de oordeelpartikels.

Het aandeel dat de directe rede heeft in de krant, neemt echter toe. Bovendien stijgt het aantal subjectieve elementen binnen de directe rede. Het aandeel van de deelcitaten neemt af, maar het gaat daarbij in beide periodes om minder dan 2% van het totale aantal woorden in het deelcorpus.

Wat betekent dit nu voor de conversationalisatietheorie? Treedt er in Nederlandse kranten conversationalisatie op of niet? De resultaten leveren geen overtuigend bewijs. Het gebruik van de directe rede neemt wel significant toe en de directe rede wordt bovendien subjectiever. Het antwoord op de vraag is volgens ons dus een genuanceerd 'ja'. Er treedt geen conversationalisatie op in de zin die de literatuur doet verwachten, namelijk door het simpelweg overnemen van subjectiviteitskenmerken; in plaats daarvan wordt conversatie op een andere manier in de tekst gebracht, namelijk door het letterlijk opnemen van conversatie in de vorm van de directe rede. Met andere woorden: de bronnen converseren in de journalistieke tekst en brengen zo hun eigen, personagesubjectieve elementen in.

De toename van het gebruik van de directe rede in de Nederlandse kranten is ook te interpreteren als een uitdrukking van veranderingen in de professionele opvattingen binnen de journalistiek. Het betreft dan een veranderde opvatting over de scheiding

tussen feiten en meningen (Sanders, 2009). Dit punt behoeft nader onderzoek, onder andere in stijlgidsen van de verschillende, hier onderzochte kranten.

Tot nog toe waren de vergelijkingen in dit onderzoek enigszins ongenueanceerd: alle artikelen uit een jaar zijn samengenomen en vergeleken met alle artikelen uit het andere jaar, zonder daarbij de verschillende genres binnen de journalistiek in aanmerking te nemen. Het is bijvoorbeeld heel wel mogelijk dat de veranderingen binnen het harde nieuws anders liggen dan die binnen de opiniërende teksten. Ook is het mogelijk dat de verschuivingen per krant verschillen. In ons huidige onderzoek is niet voldoende materiaal aanwezig om dergelijke genuanceerde vergelijkingen betrouwbaar uit te voeren.

Daarom zal het hier geanalyseerde corpus uitgebreid worden tot een corpus van ongeveer 3 miljoen woorden: 1 miljoen woorden conversatie uit het CGN, 1 miljoen woorden uit kranten uit 2002 (reeds digitaal beschikbaar) en 1 miljoen woorden uit kranten uit 1950 (moet digitaal beschikbaar gemaakt worden). Dit uitgebreide corpus maakt het mogelijk vergelijkingen tussen genres en/of tussen kranten te maken. Daarbij is het, zoals blijkt uit de resultaten van het huidige onderzoek, cruciaal dat de verschillende stemmen in de teksten (spreker en personages) steeds precies gescheiden worden. Daartoe wordt onderzoek gedaan naar de mogelijkheden om directe rede en deelcitaten automatisch te herkennen in beide krantendeelcorpora.

## Noten

1. Gaarne bedanken wij twee anonieme beoordelaars voor hun waardevolle adviezen bij een eerdere versie. Alle onvolkomenheden in dit artikel komen vanzelfsprekend uitsluitend voor onze verantwoordelijkheid.
2. Vanaf nu gebruiken we alleen de term 'subjectiviteit'.
3. Hiervoor is gebruik gemaakt van Tadpole, een programma voor automatische morpho-syntactische analyse en parsering van Nederlandse teksten (Van den Bosch et al., 2007).
4. Naast deze lexicogrammaticale analyse is ook een analyse van een deel van het corpus uitgevoerd op tekstueel niveau. Hiervoor zijn de teksten geanalyseerd voor subjectieve coherentierelaties (Mann & Thompson, 1988). Deze analyse wordt besproken in Vis (in druk).
5. Zelf toegevoegd.
6. Hiervoor is de rekenmodule gebruikt die te vinden is op de website <http://ucrel.lancs.ac.uk/llwizard.html>, geraadpleegd op 9 maart 2009. Een loglikelihood-toets maakt het mogelijk om frequenties in corpora te vergelijken ook als de bestudeerde verschijnselen relatief zeldzaam zijn in één of meer van de corpora.
7. Voornaamwoord 1e persoon: 1950: 42,3/10000 woorden, 2002: 50,9/10000 woorden,  $G2(1)=3,37$ ,  $p=0,07$ ; voornaamwoord 2e persoon: 1950: 13,2/10000 woorden, 2002: 14,1/10000 woorden,  $G2(1)=0,12$ ,  $p = 0,73$ ; cognitiewerkwoorden: 1950: 58,1/10000 woorden, 2002: 59,4/10000 woorden,  $G2(1)=0,06$ ,  $p=0,81$ ; intensifiers: 1950: 150,0/10000 woorden, 2002: 157,8/10000 woorden,  $G2(1)=0,83$ ,  $p=0,36$ .
8. Voornaamwoorden 1e persoon: 1950: 32,2/10000, 2002: 6,2/10000,  $G2(1)= 87,76$ ,  $p<.01$ ; voornaamwoorden 2e persoon: 1950: 9,5/10000, 2002: 3,3/10000,  $G2(1)= 14,12$ ,  $p<.01$ ; modale bijwoorden: 1950: 31,3/10000, 2002: 22,4/10000,  $G2(1)= 6,39$ ,  $p<.05$ ; intensifiers: 1950: 135,7/10000, 2002: 115,6/10000,  $G2(1)=6,87$ ,  $p<.05$ .
9. Modale werkwoorden: 1950: 133,7/10000, 2002: 99,5/10000,  $G2(1)=21,65$ ,  $p<.01$ ; deiktische elementen: 1950: 27,6/10000, 2002: 16,6/10000,  $G2(1)=11,91$ ,  $p<.01$ .
10. 1950: 69,8/10000, 2002: 92,5/10000,  $G2(1)=13,37$ ,  $p<.01$ .

## Literatuur

- Askehave I. & Swales J.M. (2001). Genre identification and communicative purpose: a problem and a possible solution. *Applied Linguistics*, 22(2), 195-212.
- Bekker, B. (2006). *De feiten verdraaid. Over tekstvolgorde, talige markering en sprekerbetrokkenheid*. Tilburg: Dissertatie Universiteit van Tilburg.
- Bhatia, V. (2005). Generic patterns in promotional discourse. In H. Halmari & T. Virtanen (Eds.), *Persuasion across Genre* (pp. 213-225). Amsterdam: John Benjamins.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. (2004). Historical Patterns for the Grammatical Marking of Stance: A Cross-Register Comparison, *Journal of Historical Pragmatics*, 5 (1), 107-136.
- Biber, D., Johansson, S., Leech G., Conrad S. & Finegan E. (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Bosch, A. van den, Busser, G.J., Daelemans, W., & Canisius, S. (2007). An efficient memory-based morphosyntactic tagger and parser for Dutch. In F. van Eynde, P. Dirix, I. Schuurman & V. Vandeghinste (Eds.), *Selected Papers of the 17th Computational Linguistics in the Netherlands Meeting, Leuven, Belgium*, (pp. 99-114).
- Clark, H.(1996). *Using language*. Cambridge: Cambridge University Press.
- Cotter, C. (2003). Prescription and practice. Motivations behind change in news discourse. *Journal of Historical Pragmatics*, 4(1), 45-74.
- Fairclough, N. (1992). *Discourse and Social Change*. Cambridge: Polity.
- Fairclough, N. (1994). Conversationalization of public discourse and the authority of the consumer. In R. Keat, N. Whiteley & N. Abercrombie (Eds.), *The Authority of the Consumer* (pp. 253-268). London: Routledge.
- Fairclough, N. & Mauranen, A. (1997). The conversationalisation of political discourse. *Belgian Journal of Linguistics*, 11, 89-119.
- Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J. & van den Toorn, M.C. (1997). *ANS. Algemene Nederlandse Spraakkunst* Tweede geheel herziene editie. Groningen: Martinus Nijhoff.
- Langacker, R. (1990). Subjectification. *Cognitive Linguistics*, 1, 5-38
- Lyons, J. (1994). Subjecthood and subjectivity. In M. Yaguello (Ed.), *Subjecthood and Subjectivity: Proceedings of the Colloquium 'The Status of the Subject in Linguistic Theory'* (pp. 9-17). Paris: Ophrys.
- Mann, W. & Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3), 243-281.
- Oostdijk, N. (2000). The Spoken Dutch Corpus Project. *The ELRA Newsletter*, 5(2), 4-8.
- Pearce, M. (2005). Informalization in UK party election broadcasts: 1966-97. *Language & Literature*, 14(1), 65-90.
- Sanders, J. (2009) De verdeling van verantwoordelijkheid tussen journalist en nieuwsbron. Vorm en functie van citaatmengvormen in journalistieke genres. *Tijdschrift voor Taalbeheersing*, 31(1), 1-17.
- Sanders, J. & W. Spooren (1997). Subjectivity, perspectivization, and modality from a cognitive linguistic point of view. In W.A. Liebert, G. Redeker & L. Waugh (Eds.), *Discourse and Perspective in Cognitive Linguistics* (pp. 85-112). Amsterdam: Benjamins.

- Scheibman, J. (2002). *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation*. Amsterdam: John Benjamins.
- Steen, G. (2003). Conversationalization in discourse: Stylistic changes in editorials of *The Times* between 1950 and 2000. In L. Lagerwerf, W. Spooren & L. Degand (Eds.), *Determination of Information and Tenor in Texts: Multidisciplinary Approaches to Discourse* (pp. 115-124). Münster: Nodus Publikationen.
- Vis, K. (in druk). Using RST to analyze subjectivity in text and talk. In M. Choinski & L. Wiraszka (Eds.), *Cognitive Linguistics in Action: From Theory to Application and Back*. Berlin etc.: Mouton de Gruyter.
- Wiebe, J. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20(2), 233-287.

### Bijlage. Voorbeeld van een (deel van een) geannoteerde tekst

Oorspronkelijke tekst: *Passage uit artikel uit Algemeen Dagblad, 3 mei 2002, binnenland*

Ook bij Sevenstar wist men niets van de smokkel. ‘Dit is voor ons absoluut niet te controleren’, zegt Sevenstar-directeur Koolhof.

---

#### Analyse

---

```
<sent id="26">
<br/>
<pau ref="Abi3.txt.26" s="UNKNOWN">
<pw ref="Abi3.txt.26.1" pos="BW()" lem="ook" subj="y" subjcat="modal" subjsct="part">Ook</pw>
<pw ref="Abi3.txt.26.2" pos="VZ(init)" lem="bij">bij</pw>
<pw ref="Abi3.txt.26.3" pos="N(eigen,ev,basis,zijd,stan)" lem="Sevenstar">Sevenstar</pw>
<pw ref="Abi3.txt.26.4" pos="WW(pv,verl,ev)" lem="weten" subj="y" subjcat="verbcogn" sub-
itense="past">wist</pw>
<pw ref="Abi3.txt.26.5" pos="VNW(pers,pron,nomin,red,3p,ev,masc)" lem="men">men</pw>
<pw ref="Abi3.txt.26.6" pos="VNW(onbep,pron,stan,vol,3o,ev)" lem="niets">niets</pw>
<pw ref="Abi3.txt.26.7" pos="VZ(init)" lem="van">van</pw>
<pw ref="Abi3.txt.26.8" pos="LID(bep,stan,rest)" lem="de">de</pw>
<pw ref="Abi3.txt.26.9" pos="N(soort,ev,basis,zijd,stan)" lem="smokkel">smokkel</pw>
<pl ref="Abi3.txt.26.10" pos="LET()" lem=".">.</pl>
</pau>
</sent>
<sent id="27">
<pau ref="Abi3.txt.27" s="UNKNOWN">
<DS>
<pw ref="Abi3.txt.27.1" pos="LET()" lem="'">'</pw>
<pw ref="Abi3.txt.27.2" pos="VNW(aanw,pron,stan,vol,3o,ev)" lem="dit"><mrw ty-
pe="met">Dit</mrw></pw>
<pw ref="Abi3.txt.27.3" pos="WW(pv,tgw,ev)" lem="zijn" subitense="pres">is</pw>
<pw ref="Abi3.txt.27.4" pos="VZ(init)" lem="voor"><mrw type="met">voor</mrw></pw>
<pw ref="Abi3.txt.27.5" pos="VNW(pr,pron,obl,vol,1,mv)" lem="ons" subj="y" subjcat="1stp" subjs-
cat="pl">ons</pw>
<pw ref="Abi3.txt.27.6" pos="ADJ(vrij,basis,zonder)" lem="absoluut">absoluut</pw>
<pw ref="Abi3.txt.27.7" pos="BW()" lem="niet">niet</pw>
<pw ref="Abi3.txt.27.8" pos="VZ(init)" lem="te">te</pw>
<pw ref="Abi3.txt.27.9" pos="WW(inf,vrij,zonder)" lem="controleren">controleren</pw>
<pw ref="Abi3.txt.27.10" pos="LET()" lem="'">'</pw>
</DS>
<pw ref="Abi3.txt.27.11" pos="LET()" lem=",">,</pw>
<pw ref="Abi3.txt.27.12" pos="WW(pv,tgw,met-t)" lem="zeggen" subj="y" subjcat="verbcogn" sub-
itense="pres">zeqt</pw>
<pw ref="Abi3.txt.27.13" pos="N(soort,ev,basis,zijd,stan)" lem="sevenstar-directeur">Sevenstar-
directeur</pw>
<pw ref="Abi3.txt.27.14" pos="N(eigen,ev,basis,zijd,stan)" lem="Koolhof">Koolhof</pw>
<pl ref="Abi3.txt.27.15" pos="LET()" lem=".">.</pl>
</pau>
</sent>
```

---

#### Toelichting:

De subjectieve elementen zijn onderstreept.

Subj="y" staat voor subjectief element.

<DS> staat voor directe rede (Direct Speech).

---